

Co je statistika, proměnné, úrovně měření

POLb1139 Statistické myšlení v sociálních vědách



“Without data you’re just another person with an opinion”

- W. Edwards Deming

Už jste se někdy v minulosti učil/a statistiku?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Ano	2	33,3	33,3	33,3
	Ne	4	66,7	66,7	100,0
	Total	6	100,0	100,0	

Očekáváte, že tento kurz bude ve srovnání s jinými kurzy jednoduchý, náročný, nebo na stejné úrovni náročnosti?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Spíše jednoduchý	1	16,7	16,7	16,7
	Tak nastejno	3	50,0	50,0	66,7
	Spíše náročný	1	16,7	16,7	83,3
	Rozhodně náročný	1	16,7	16,7	100,0
	Total	6	100,0	100,0	

Očekáváte, že bude tento kurz pro Vás ve srovnání s jinými kurzy užitečný?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Spíše užitečný	1	16,7	16,7	16,7
	Rozhodně užitečný	5	83,3	83,3	100,0
	Total	6	100,0	100,0	

Co je to podle Vás statistika?

- Nástroj/technika, z jejíchž výstupů následně můžeme usuzovat a dovozovat závěry, hledat vzorce, odpovídat na výzkumné otázky,...
- Souhrn informací, ze kterých lze vyvodit nějaký úsudek
- Systém usporiadania dát pre ich operacionalizáciu
- Veda kvantifikácie dát
- Vyjadreni reality zkoumane problematiky v pouzitelnejsich ciselnych hodnotach

K čemu je podle Vás statistika užitečná?

- K zjištění stavu nějakého zkoumaného problému ve společnosti
- Ku kvantifikácii dat
- Pre (nie iba) vedeckú inferenciu dnes už prakticky v akomkoľvek odbore
- Při správném uplatnění je velmi vhodným nástrojem na zlepšení (zvýšení exaktnosti) výzkumné strategie
- Užitečnost statistiky je převážně v tom, že díky ní lze vyvodit nějaké závěry, popřípadě popisuje nějaký stav, který se převážně odchyluje od normálu apod.

Proč jste se rozhodl/a zapsat zrovna do tohoto kurzu?

- Abych se naučil úplně novou dovednost (a ne další znalost)
- Jelikož si myslím, že je to jedna z mála příležitostí, jak se rozvinout v oblasti, která se mnohdy vyučuje až nepochopitelně sofistikovaně. Zároveň cítím, že účely práce s daty nyní nebo později v pracovním procesu, je statika nezbytným nástrojem. A v neposlední řadě; kdy jindy než nyní a ještě k tomu v tak komorním prostředí devíti mužů :D
- Praktické poznatky ku bakalářské práci
- Protože ve většině studií sledávám určité experimenty spjaté se statistikou - s mým neuspokojivým přehledem ve statistice jim však moc nerozumím.
- Z důvodu naučení se nových znalostí a praktických dovedností, které využiju při psaní bakalářské práci

Statistika

- *„It is a key component in the scientific toolbox and one of the main ways we have of describing the natural world and of finding out how it works.“ (Hector 2021)*
- *„Statistics is a bit like sticking your finger into a revolving fan blade: sometimes it’s very painful, but it does give you the power to answer interesting questions.“ (Field 2009)*

Statistika

- Souhrn technik, pomocí kterých se realizuje komplexní práce s daty:
 - Sběr
 - Analýza
 - Interpretace
 - *(a vizualizace)*
- Přesahuje jednotlivé tematické oblasti vědy
- Hledá odpovědi na množství otázek
- Dívá se na svět kvantitativním pohledem

Jak to funguje (Spiegelhalter 2019)

- Problem – Plan – Data – Analysis - Conclusion
1. Na základě pozorování si stanovíme problém
 2. Naplánujeme si postup
 3. Stanovíme si teorii vysvětlující náš problém (hypotézy)
 4. Získáme vhodná data a otestujeme platnost hypotézy
 5. Vytvoříme důvěryhodný závěr z našich analýz

Jak to funguje (Spiegelhalter 2019)

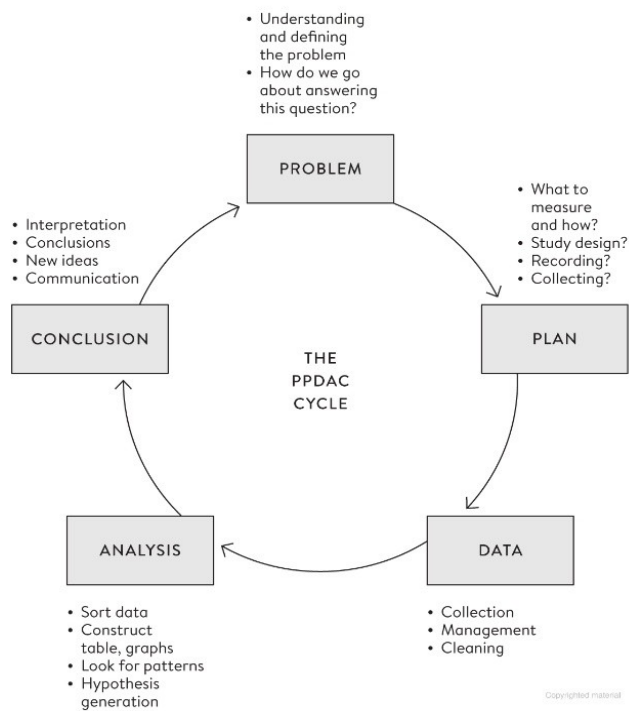
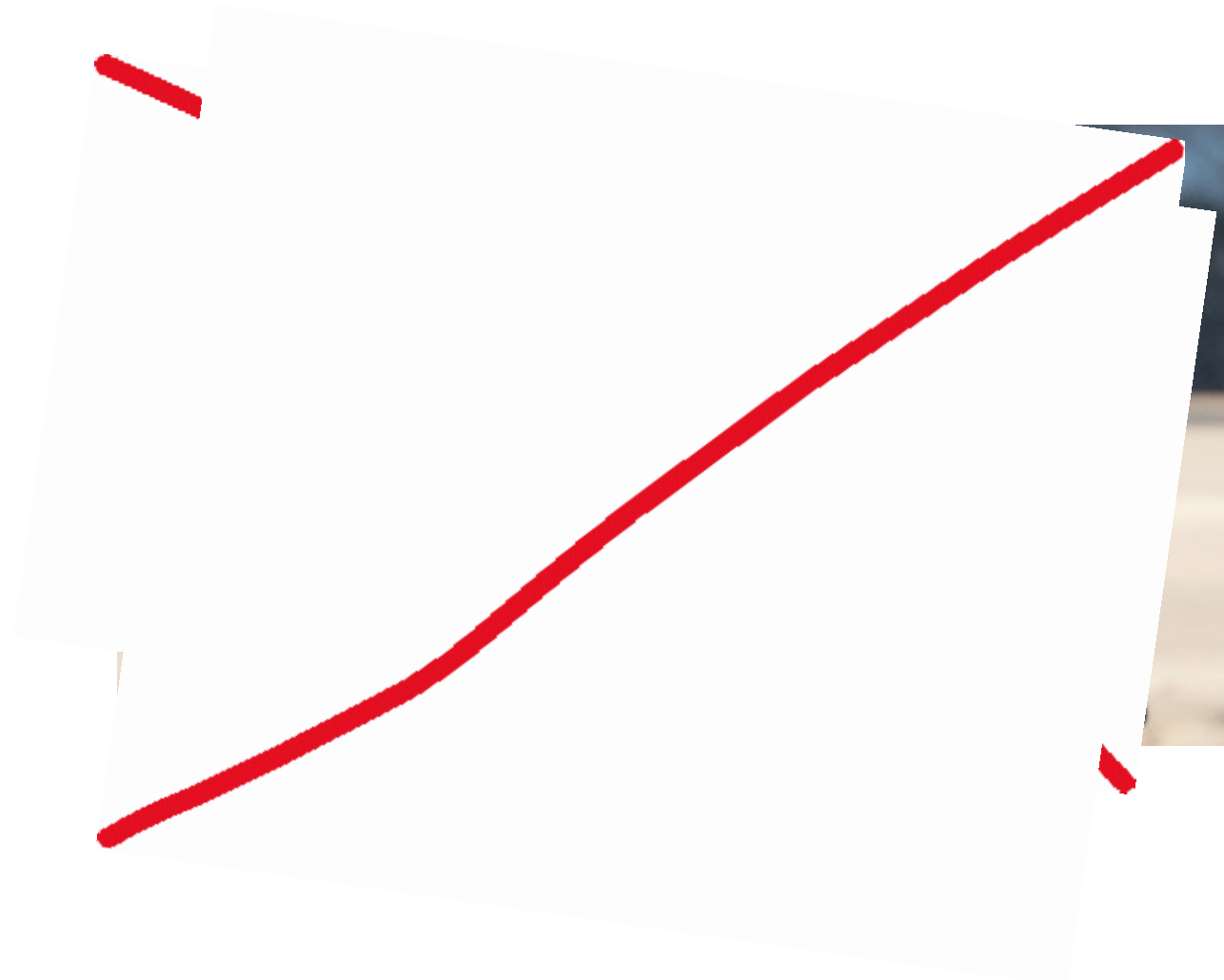


Figure 0.3

The PPDAC problem-solving cycle, going from Problem, Plan, Data, Analysis to Conclusion and communication, and starting again on another cycle.



Data: proměnné a případy

- Statistická analýza bez (velkého počtu) dat není možná
- Potřebujeme vědět:
 - Co měříme (koncept)
 - Jak to měříme (operacionalizace)
 - Na jaké úrovni měříme (jednotlivci, regiony, státy)
 - Co je jednotka analýzy (volič, politická strana, obec)

Rychlost auta

- Jak měřit: maximální rychlost auta v km/hod (operacionalizace)
 - Úroveň měření: individuální
 - Jednotka analýzy: auto
-
- Jednoduché a vcelku nesporné



Spokojenost s demokracií

- Jak měřit?
- Úroveň měření? (občané vs. státy)
- Občané:
 - Jak spokojeni jste se stavem demokracie ve vaší zemi? (škála 1-10)
 - Souhlasíte s následujícím tvrzením? Liberální demokracie s pravidelnými volbami je vhodnější než vláda silného lídra. (ano/ne)
- Odlišné způsoby měření → odlišné proměnné

Proč to nemíchat

- Úroveň měření vs. interpretace dosažených zjištění
- Ekologická chyba – aplikace výsledků na určitou úroveň bez toho, aby k tomu opravňovala podoba dat a analýzy
- Příklad:
 - Zjištění: konzervativní strany mají vyšší podporu v obcích s vyšší mírou nezaměstnanosti
 - Interpretace: nezaměstnaní lidé volí konzervativní strany
 - **Kde je chyba?**

Proměnná

- Způsob uchopení jevů, které zkoumáme a měříme
- Její hodnota variuje (jinak jde o konstantu)
- Klíčový prvek nejen pro kvantitativní výzkum
- Denní teplota, počet hlasu strany ve volbách, cena ovoce, e-mailové adresy zaměstnanců univerzity

Dělení proměnných

- Dle postavení ve výzkumu:
 - Nezávislá (nezávisle) proměnná – předpokládaná příčina
 - Závislá (závisle) proměnná – předpokládaný následek
- Předpoklad kauzálního vztahu NP → ZP
- Základ pro formulaci hypotéz
- Přítomnost třetích proměnných a jejich kontrola

Dělení proměnných dle úrovně měření

		Možnost seřazení hodnot	Stejně rozdíly hodnot	Věcně významná nula
Kategorické	Nominální	-	-	-
	Ordinální	+	-	-
Kardinální	Intervalové	+	+	-
	Poměrové	+	+	+

- Kardinální:
 - Diskrétní – pouze celá čísla
 - Spojitá – teoreticky nekonečné množství hodnot

Příklady proměnných

- Nominální:
 - jména osob, názvy obcí
 - pouze dvě hodnoty – binární (dummy) – student (ano/ne), účast ve volbách (ano/ne)
- Ordinální – stupně vzdělání, vojenské hodnosti
- Intervalová – teplota v st. Celsia, poschodí v domě
- Poměrové – počet členů domácnosti, míra nezaměstnanosti
- **Důležitý je obsah, ne formální zápis** – jména osob se dají seřadit abecedně, ale nejde o věcně smysluplné pořadí (stejně tak telefonní čísla, adresy...)

Klíčové pojmy na závěr

- Reliabilita:

- Schopnost dosažení stejných výsledků při opakovaném měření
- Jak spolehlivá je naše operacionalizace a naše měření?
- Bude stejná hodnota naměřena zítra?
- Když kategorizací a měřením pověřím dva různé lidi, naměří stejnou hodnotu?

- Validita:

- Schopnost měření toho, co skutečně chceme měřit
- Schopnost generalizace našich zjištění (externí validita) – jiná přednáška

Validita a reliabilita (Lundberg 2006)

